BUSSTEPP Homework 1:
A Tutorial on Numerical Methods with Quantum Mechanics

August 29, 2003

Abstract

In numerical lattice QCD, one evaluates the path integral with a Monte Carlo method. This tutorial introduces novices to the techniques using quantum mechanics for illustration. The examples used are (i) the harmonic oscillator, (ii) an anharmonic oscillator with $V(x) = \frac{1}{2}m\omega^2x^2 + \frac{1}{4}\lambda x^4$, and a "double hump" with $V(x) = -\frac{1}{2}m\omega^2x^2 + \frac{1}{4}\lambda x^4$.

Part of the homework is to write a computer program. Students with little programming experience may team up with those who do. You don't need more knowledge than compiling and running a simple program, and a random number generator, which you can take from a standard library.

1 Review of Path Integrals

Let us recall Feynman's derivation of the path integral formulation of quantum mechanics. This is in many textbooks, so we shall be sketchy here. We shall set $\hbar = 1$ and $c = 1$, so that p, m, ω, E, $1/t$, and $1/x$, have the same dimension (as usual in particle physics).

The Hamiltonian is

$$H = \frac{p^2}{2m} + V(x). \tag{1}$$

We would like to derive an expression for the propagator, *i.e.*, the amplitude that the particle propagates from $x = x_0$ at $t = 0$ to $x = x_T$ at $t = T$. We have

$$\langle x(T)|x(0)\rangle = \langle x_T|e^{i\hat{H}T}|x_0\rangle = \sum_n \langle x_T|n\rangle e^{iE_nT}\langle n|x_0\rangle = \sum_n \psi_n^*(x_T)\psi_n(x_0)e^{iE_nT}, \tag{2}$$

where the n labels the eigenstates of $\hat{H}$ with eigenvalue E_n and wave functions $\psi_n(x)$. The last two expressions remind us that a lot of information is encoded in the propagator.

Divide the total time interval into N small intervals, for convenience of equal size $T/N = \delta$. At the end of each interval, insert

$$1 = \int dx_i |x_i\rangle\langle x_i|, \quad i = 1, \ldots, N-1. \tag{3}$$

Then

$$\langle x_T|e^{-i\hat{H}T}|x_0\rangle = \int \prod_{i=1}^{N-1} dx_i \prod_{i=0}^{N-1} \langle x_{i+1}|e^{-i\hat{H}\delta}|x_i\rangle, \tag{4}$$

where $x_N = x_T$. Note that there are $N-1$ integrations and N matrix elements.

Now *approximate*

$$\langle x_{i+1}|e^{-i\hat{H}\delta}|x_i\rangle \approx \langle x_{i+1}|e^{-i\hat{V}(x)\delta/2}e^{-i\hat{p}^2\delta/2m}e^{-i\hat{V}(x)\delta/2}|x_i\rangle, \tag{5}$$

and re-write the right-hand side

$$\langle x_{i+1}|e^{-i\hat{V}(x)\delta/2}e^{-i\hat{p}^2\delta/2m}e^{-i\hat{V}(x)\delta/2}|x_i\rangle = e^{-iV(x_{i+1})\delta/2}\langle x_{i+1}|e^{-i\hat{p}^2\delta/2m}|x_i\rangle e^{-iV(x_i)\delta/2}. \tag{6}$$

Exercise I.1: Show that

$$\langle x_{i+1}|e^{-\hat{p}^2 a/2m}|x_i\rangle = \sqrt{\frac{m}{2\pi a}}e^{-m(x_{i+1}-x_i)^2/2a}. \tag{7}$$

Note that here the time interval is imaginary: $\delta = -ia$, with a real. Why?

Combining Eq. 7 with Eqs. (4) and (6), one has (for imaginary time, $T \to -iT$)

$$\langle x_T|e^{-\hat{H}T}|x_0\rangle = \lim_{N\to\infty} \int \mathcal{D}x \ \exp\left(a\sum_{i=0}^{N-1} L_i\right), \qquad \mathcal{D}x = \prod_{i=1}^{N-1} dx_i \sqrt{\frac{m}{2\pi a}}, \tag{8}$$

where the limit is taken with T fixed, $a = T/N$. The (discrete time) Lagrangian is

$$L_i = -\tfrac{1}{2}m\left(\frac{x_{i+1}-x_i}{a}\right)^2 - \tfrac{1}{2}V(x_{i+1}) - \tfrac{1}{2}V(x_i), \tag{9}$$

which one recognizes as a discrete approximation to the kinetic energy and the average of the potential energy over two times.

Because of the sum over i in Eq. (8), we could have also defined L_i with $V(x_i)$ alone, instead of the average.

We usually set $S = -a\sum_i L_i$ and call S the Euclidean action. (In field theory, "imaginary time" has metric $(+,+,+,+)$.) For a sensible theory, S must be bounded below.

In this tutorial, we wish to evaluate path integrals such as the right-hand side of Eq. (8) numerically. Therefore we maintain the imaginary-time formalism. Then the integrals die off rapidly for large x. With real time, there would be wild oscillations, and (particularly in field theory) drastic cancellations between adjacent patches of configuration space $\{x_i : i =$

$0, N-1\}$. That is unfeasible numerically. As we shall see below, it is easy to get at most desired information. (And in quantum mechanics, all desired information.)

You can obtain the ground-state energy and (squared) wave function as follows. Set $x_0 = x_T = x$. Then

$$\langle x|e^{-\hat{H}T}|x\rangle = \sum_n |\langle x|n\rangle|^2 e^{-E_n T} \approx |\langle x|0\rangle|^2 e^{-E_0 T}, \tag{10}$$

the last for large T. Evaluating the right-hand side of Eq. (8) numerically as a function of T and x, for $E_0 a = E_0 T/N \ll 1$, yields the desired information. For more details, consult Lepage, "Lattice QCD for Novices."

In field theory, the ground state is the "vacuum," and we are usually more interested in excited states, which correspond to (elementary or composite) particles. To this end it is convenient to set $x_T = x_0$ and integrate over x_0 too.

$$Z = \int \mathcal{D}x \, \exp(-S), \qquad \mathcal{D}x = \prod_{i=0}^{N-1} dx_i \sqrt{\frac{m}{2\pi a}}, \tag{11}$$

$$Z[J] = \int \mathcal{D}x \, \exp\left(-S + \sum_i J_i x_i\right). \tag{12}$$

A very useful object is the *connected* 2-point correlation function:

$$\frac{d^2 \ln Z[J]}{dJ_j dJ_0} = \frac{1}{Z}\int \mathcal{D}x \, x_j x_0 \, e^{-S} - \frac{1}{Z}\int \mathcal{D}x \, x_j \, e^{-S} \frac{1}{Z}\int \mathcal{D}x \, x_0 \, e^{-S}. \tag{13}$$

This can be generalized to n-point functions and also to correlation functions of $f(x_i) \mapsto f(d/dJ)$.

Exercise I.2: Show that the integrals on the right-hand side can be expressed as matrix elements and exponentials $e^{-E_n t}$, *viz.*

$$\langle x_j \rangle := \frac{1}{Z}\int \mathcal{D}x \, x_j \, e^{-S} \quad \overset{\text{large } T}{\to} \quad \langle 0|\hat{x}|0\rangle \tag{14}$$

$$\langle x_j x_0 \rangle_c := \frac{d^2 \ln Z[J]}{dJ_j dJ_0} \quad \overset{\text{large } T}{\to} \quad \sum_{n\neq 0} |\langle 0|\hat{x}|n\rangle|^2 \left[e^{-(E_n - E_0)ja} + e^{-(E_n - E_0)(T-ja)}\right] \tag{15}$$

$$= \langle x_j x_0 \rangle - \langle x_j \rangle \langle x_0 \rangle. \tag{16}$$

The notation $\langle \Phi(\{x_i\}) \rangle$ for the path-integral average should not be confused with Dirac notation for quantum mechanical matrix elements. Both are common usage.

The connected 2-point correlation functions are similar to time-ordered products. The first series dominates when $T \gg ja$, the second when $T \gg T - ja$. Since we have periodic boundary conditions in (imaginary) time, the time step $T - ja$ comes before $t = 0$.

If $T \gg ja \gg (E_2 - E_1)^{-1}$, then the first term dominates

$$\frac{d^2 \ln Z[J]}{dJ_j dJ_0} \to |\langle 0|\hat{x}|1\rangle|^2 e^{-(E_1 - E_0)ja}. \tag{17}$$

So from the connected 2-point function and a fit to the j dependence, one can obtain $E_1 - E_0$, which in a field theory would be called the mass gap. If we can find a function $f(x)$ so that $\langle 0|f(\hat{x})|1\rangle = 0$, then one can also find the mass gap of higher-lying states. This is often possible by exploiting symmetries. There are also ways to devise $\langle 0|f(\hat{x})|1\rangle \approx 0$.

2 Monte Carlo Integration

The imaginary time path integral consists of an ordinary N-dimensional integral, so it should be possible to carry out the integration numerically. For quantum mechanics with modest N, general-purpose integration routines are enough. Here, however, we introduce a technique that works for large N and, hence, also for field theory (which has a handful of integration variables for each space-time point, of order N^4).

The first ingredient is Monte Carlo integration. The idea is to choose a random configuration of the x_i, which we can denote $\{x_i\}^{(c)}$, and we can generate C such configurations. Then

$$Z = \int \mathcal{D}x\, e^{-S} = \lim_{C\to\infty} \left(\frac{m}{2\pi a}\right)^{N/2} \sum_{c=0}^{C-1} \exp\left[-S\left(\{x_i\}^{(c)}\right)\right], \tag{18}$$

$$\int \mathcal{D}x\, f(\{x_i\}) e^{-S} = \lim_{C\to\infty} \left(\frac{m}{2\pi a}\right)^{N/2} \sum_{c=0}^{C-1} f(\{x_i\}^{(c)}) \exp\left[-S\left(\{x_i\}^{(c)}\right)\right]. \tag{19}$$

An estimate of the left-hand side is achieved for C finite.

This is hopelessly inefficient for large N, because S is an extensive quantity (S grows with N). Therefore, configurations with large action make almost no contribution to the sums. They are a waste of (computer) time.

The remedy is called importance sampling. Instead of choosing all configuration with equal weight, choose them with weight e^{-S}. This is possible because S is real and bounded below. Then

$$\frac{1}{Z} \int \mathcal{D}x\, f(\{x_i\}) e^{-S} = \lim_{C\to\infty} \frac{1}{C} \sum_{c=0}^{C-1} f(\{x_i\}^{(c)}). \tag{20}$$

Again an estimate of the left-hand side is achieved for C finite.

A simple way to generate configurations is the Metropolis algorithm. A flow chart is given in Fig. 1, customized to quantum mechanics with one degree of freedom. Here `rand()` is a random number generator that returns a real number $\xi \in [0, 1)$. In lattice gauge theory, the gauge

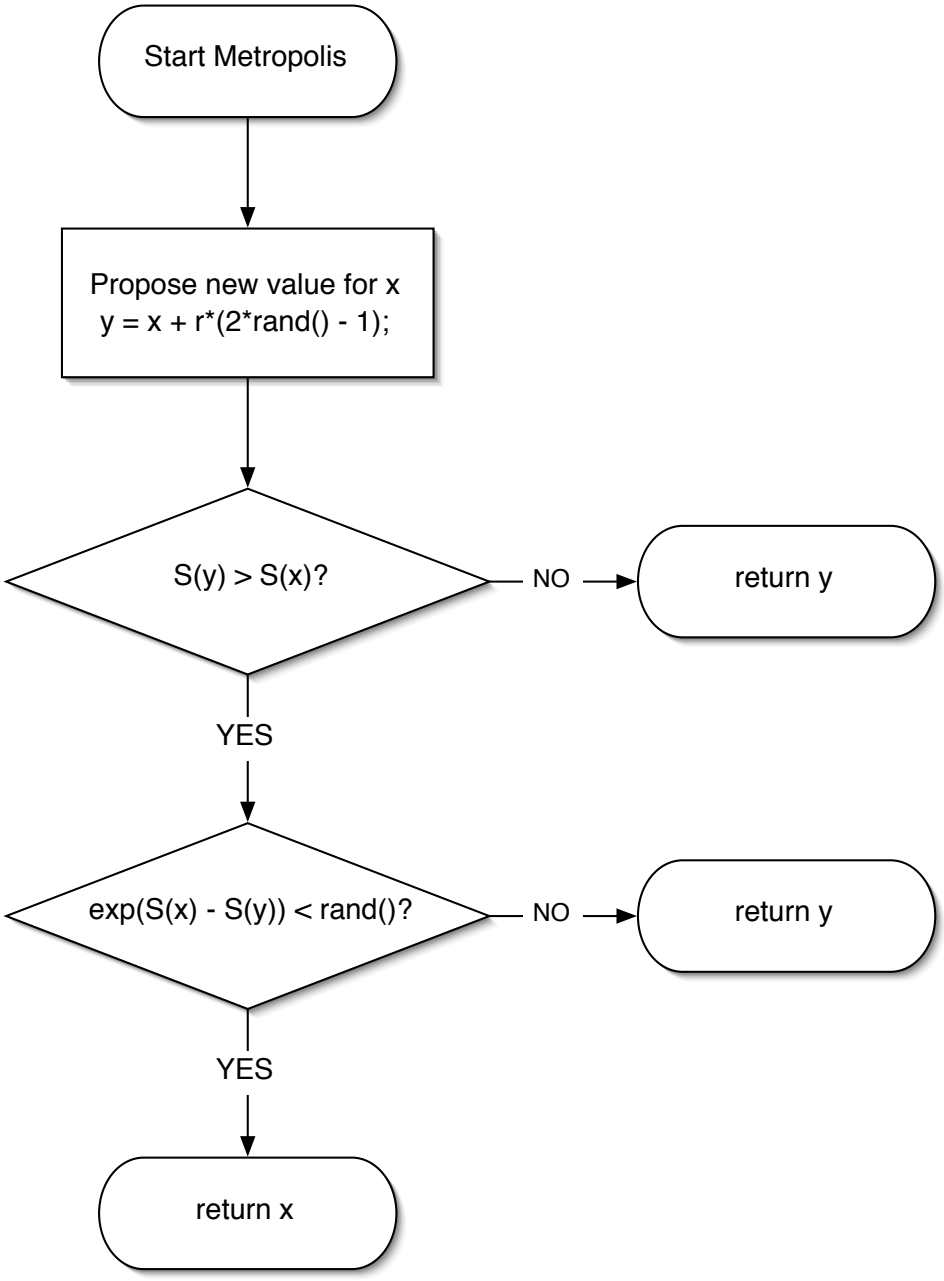


Figure 1: Flow chart of the Metropolis algorithm. For a more complicated set of degrees of freedom, the only change is to customized the proposed update from $\texttt{y} = \texttt{x} + \texttt{r} * (\texttt{2} * \texttt{rand}() - \texttt{1})$ to something appropriate to the space in which the degrees of freedom live. Here $S(x)$ need only be that part of the action that depends on the variable(s) being updated.

variables are not real numbers (they take values in a Lie group), so the proposed update has to be customized accordingly.

The parameter `r` is chosen for efficiency, so that the new values (the branches on the right in Fig. 1) are accepted around 50% (say, 40–60%) of the time. If `r` is too large, the proposed updated action is usually large and usually rejected, so the new configuration rarely changes. If `r` is too small, the new configuration is likely to be accepted, but it is hardly different from its predecessor. To move through configuration space efficiently, one must avoid these extremes.

The disadvantage of the Metropolis algorithm is that it does not move through configuration space very quickly. So the configurations used in Eq. (20) are not those from every update, but rather after repeating the Metropolis procedure often enough to obtain a statistically independent configuration. There are two simple tricks to speed up the generation of a statistically independent configuration. First, visit each variable x_i sequentially, rather than generating new values for all N variables in $\{x_i\}$ and applying the test to the new configuration:

```
for(i=0; i<N; i++) {
    x[i] = metropolis(x[i], r, S);
}
```

Second, while at any given site, repeat the procedure `Hits` times:

```
for(i=0; i<N; i++) {
    for(j=0; j<Hits; j++) x[i] = metropolis(x[i], r, S);
}
```

(For lattice gauge theory it is more efficient to move the loop over hits into the Metropolis routine, in such a way that most of the computation of S is re-used for each hit.)

Exercise I.3: Write a Metropolis updating program for the one-dimensional harmonic oscillator

$$V(x) = \tfrac{1}{2} m\omega^2 x^2 \tag{21}$$

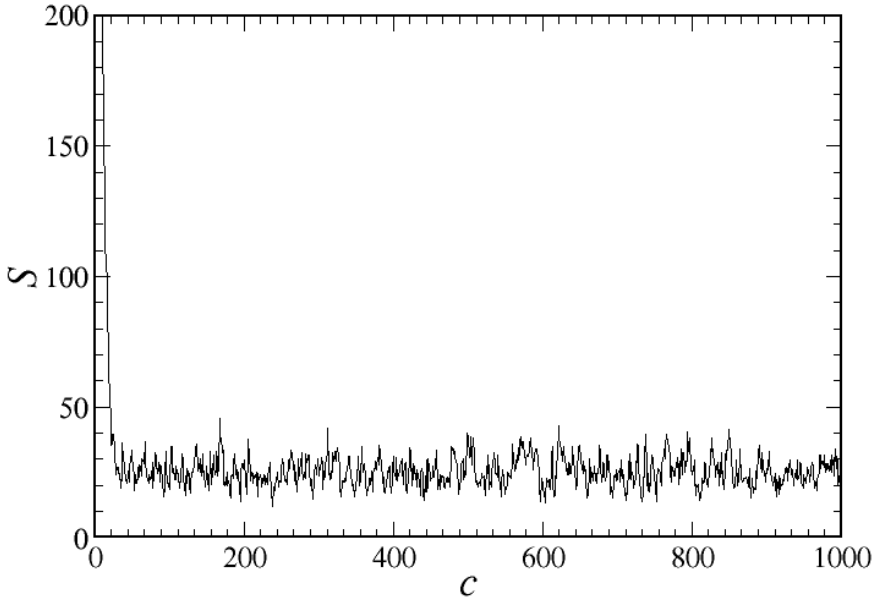
$$S = ma \sum_{i=0}^{N-1} \tfrac{1}{2}[x_{i+1} - x_i)/a]^2 + \tfrac{1}{2}(\omega a)^2 (x_i/a)^2 \tag{22}$$

The program will be easiest in lattice units, with dimensionless parameters ma and ωa, and dimensionless variables x_i/a.

Plot S and $x_{\rm avg} = \sum_{t=0}^{N-1} x_t/N$ vs. the configuration label c to check that the program is behaving sensibly. Results are in Fig. 2, using $ma = 1$, $\omega a = 1$; starting from $x_t/a = 5 \forall t$; and updating with $\mathtt{r} = 3$ and $\mathtt{Hits} = 1$. The fraction of updates accepted was 54%.

Study how these plots change when `r` is varied.

No matter how `r` is chosen, a configuration is still has too much in common with its predecessor. Furthermore, the first several configurations have too strong a memory of the (arbitrary)



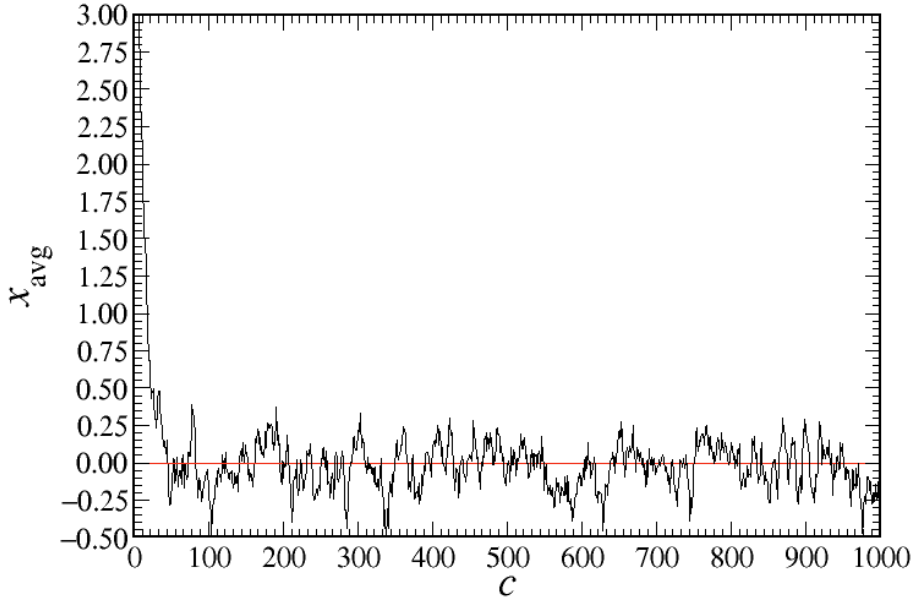


Figure 2: S and x_{avg} vs. c

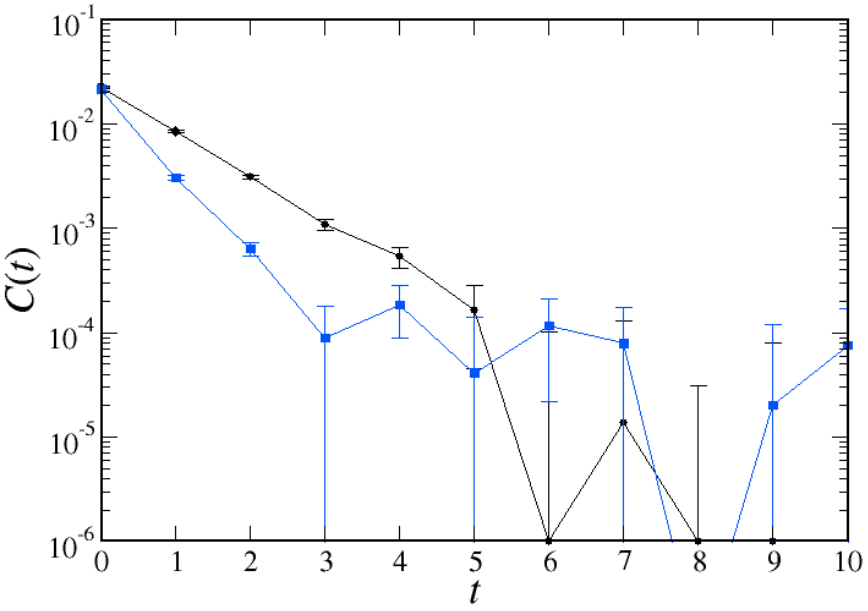


Figure 3: S and $x_{\rm avg}$ vs. c

initial configuration. These related features can be worked around by discarding the first `Ntherm` configurations (called thermalization), and then saving only every `Ndecorr`-th configuration for the ensemble average (called decorrelation). From Fig. 2 it seems that `Ndecorr` $= 20$ and `Ntherm` $= 200$ should be sufficient.

Exercise I.4: Run with the same parameters as before, except use the recommended thermalization and decorrelation. Compute the correlation functions $C_1(t) = \langle x_t x_0 \rangle$ and $C_2(t) = \langle x_t^2 x_0^2 \rangle_c$, and extract the first two energy levels. Note that the $\langle x_t x_0 \rangle$ is automatically connected, because of parity $V(-x) = V(x)$). Hence $\langle x \rangle = 0$. Furthermore, parity also guarantees that $\langle 1|\hat{x}^2|0\rangle = 0$; consequently $C_2(t)$ gives the energy of the second excited state. Verify that the energy levels are as expected, $E_1 = \omega a$, $E_2 = 2\omega a$. Results for the t dependence are in Fig. 3.
